Signals and Systems: Course Project

Author: *LIU Qiaoan 520030910220*

Course: *Spring 2022, ICE2605: Signals and Systems*
Date: *May 9, 2022*

In ICE2605, I learned the basic knowledge about signal processing. In this project, I do experiment by hand and understand what I learned in class more deeply.

Part 1

In this part, I pick a short audio clip, which is *Spring* of Joe Hisaishi. Then I use MATLAB to write the program to analyze it.

(1). First, I use `audioread` to read the wav file. From the result of `audioread`, we can get a matrix `x_N` and the sample rate `Fs`. Then we display it.

Second, I use `interp1` to do linear interpolation, which makes $x[n]$ to be a "continuous" time signal $x(t)$. Then I plot the waveform of $x(t)$.

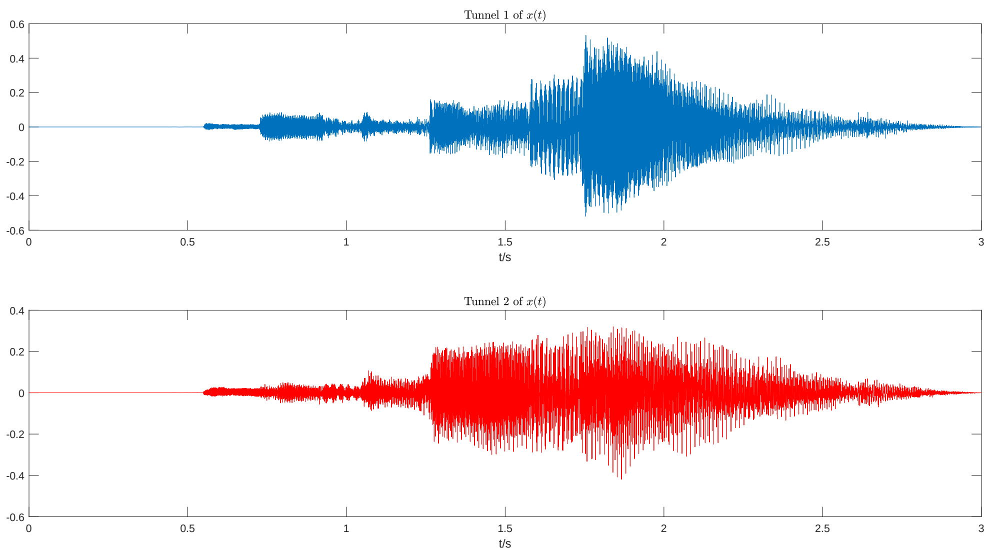


Figure 1: The waveform of $x(t)$

(2). Then, I change the x-coordinates to plot $x(t)$, $x(2t)$ and $x(t/2)$.

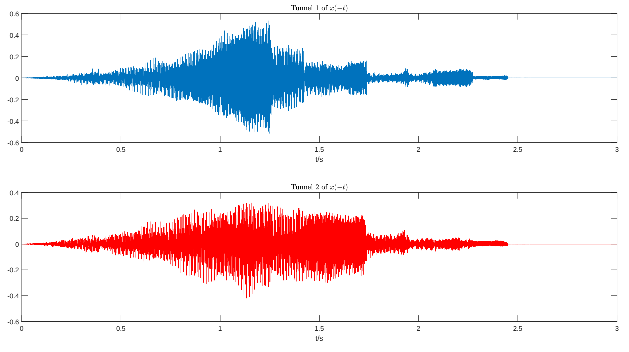


Figure 2: The waveform of $x(-t)$

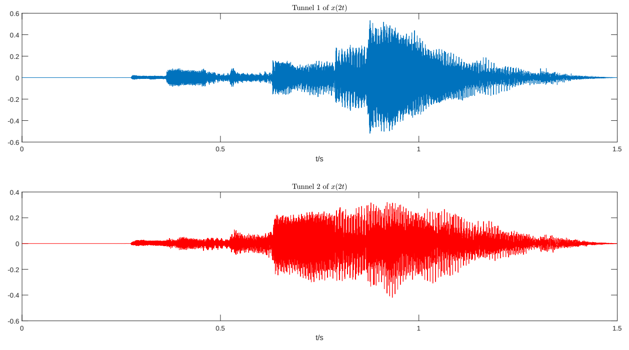


Figure 3: The waveform of $x(2t)$

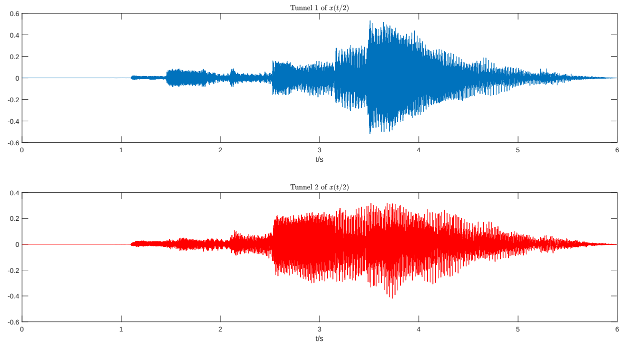


Figure 4: The waveform of $x(t/2)$

(3). In this section, I use **`fft`** to get the Discrete Fourier Transform. And the spectrum of $x(t)$ is scaled of the spectrum of $x[n]$ in frequency domain by $\frac{1}{T}$ (T is the sampling period). Then we plot the spectra of $x(t)$, $x(t)$, $x(2t)$ and $x(t/2)$. (The figrue above is a global one and the figure below is a local one from $[-5000Hz, 5000Hz]$.)

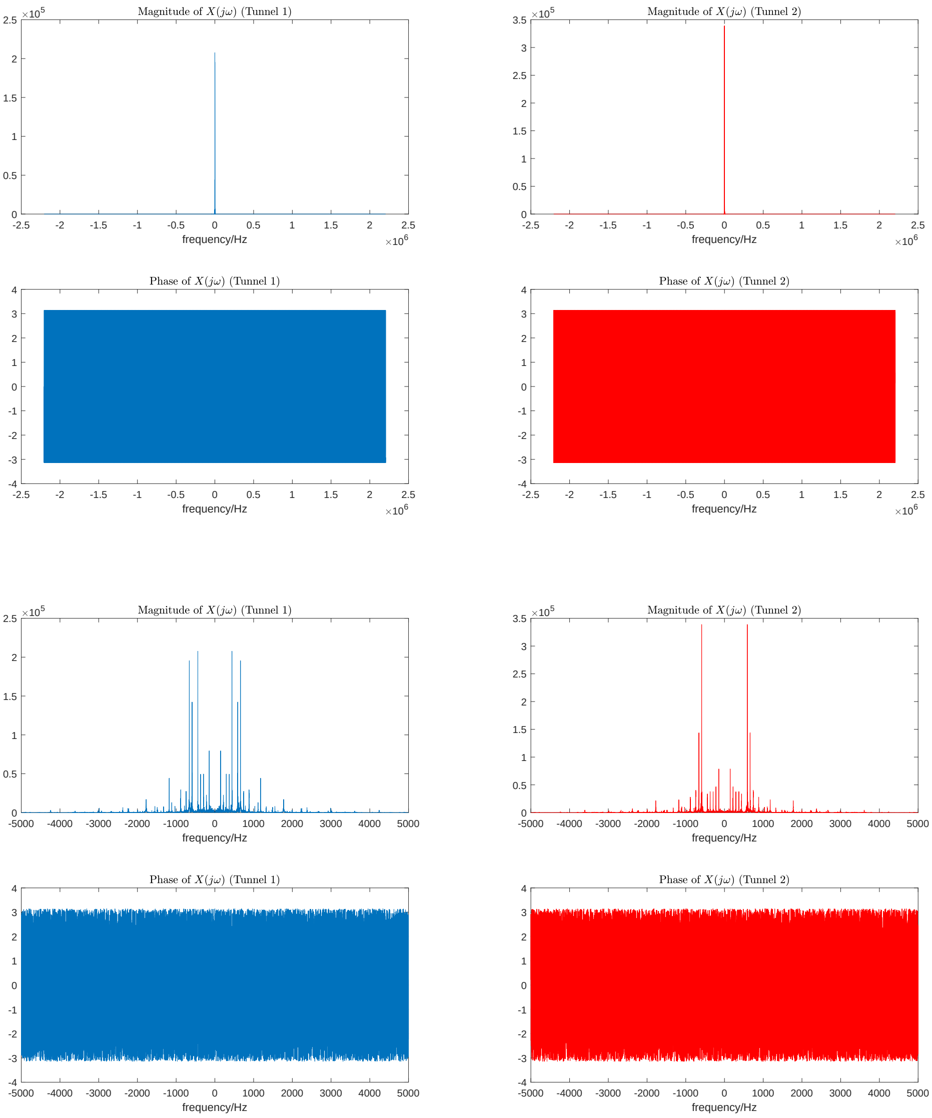


Figure 5: The spectrum of $x(t)$

We can see that the spectrum of $x(-t)$ is exactly the reversal of $x(t)$ in frequency domain. That's because $x(-t)$ is time reversal of $x(t)$, from class, I learned that $\mathcal{F}\{Rx\} = R\mathcal{F}\{x\}$. See Figure 5(a), 5(b), 6(a), 6(b).

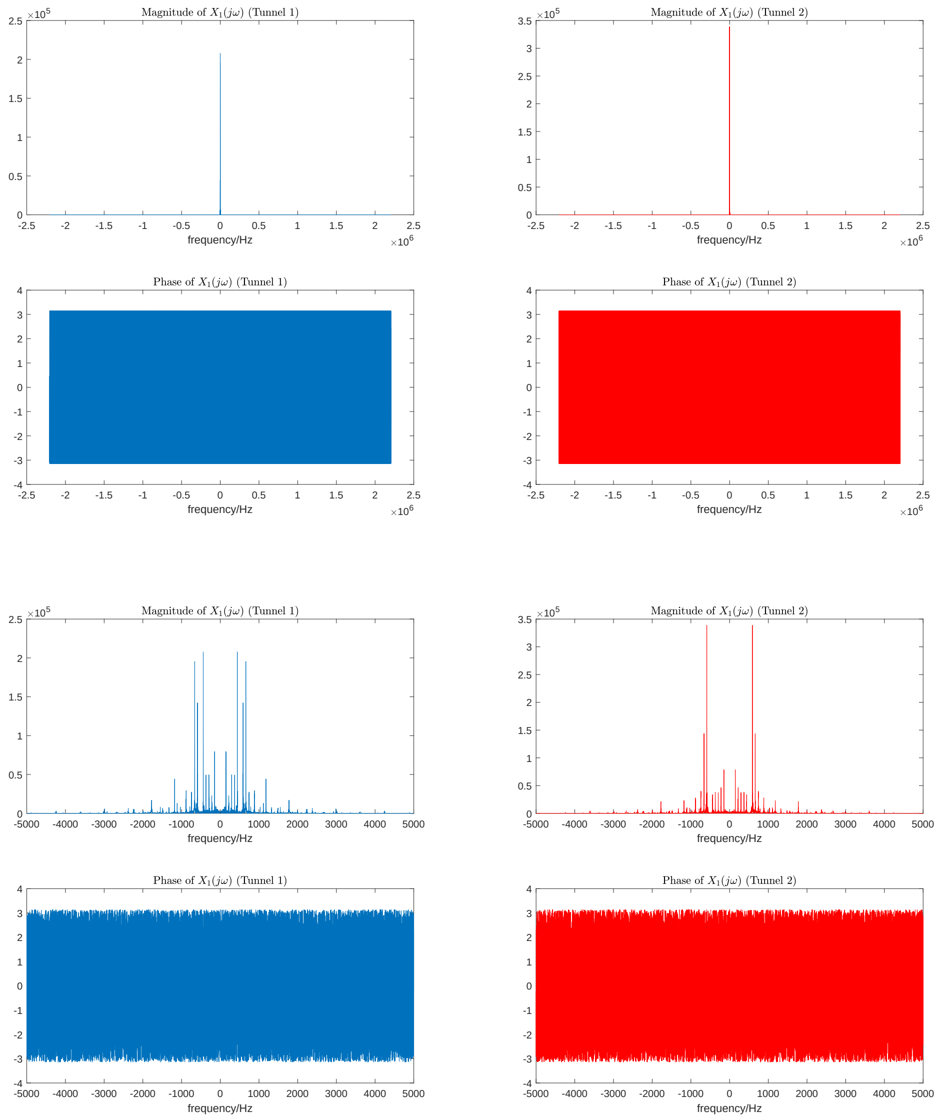
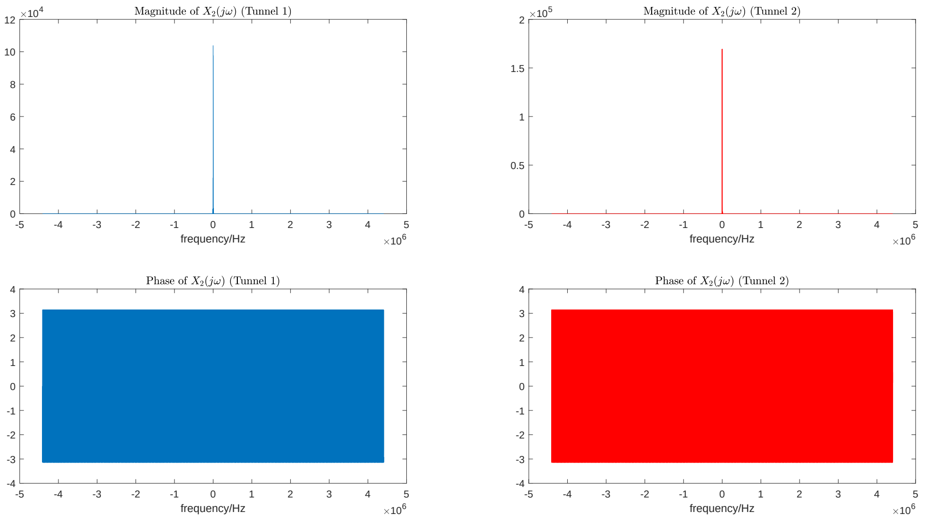


Figure 6: The spectrum of $x(-t)$

The spectrum of $x(2t)$ and $x(t/2)$ are time scaling of $x(t)$, and $x(at) \stackrel{\mathcal{F}}{\longleftrightarrow} \frac{1}{|a|}X(\frac{j\omega}{a})$. Thus, we can see Figure 5(a), 5(b), 7(a), 7(b), 8(a), 8(b). An intuitive consideration is that $x(2t)$ is a compression in time domain, then it must contain more high frequency components than $x(t)$. Similarly, $x(t/2)$ is an expension in time domain, which means it must contain more low frequency components than $x(t)$.



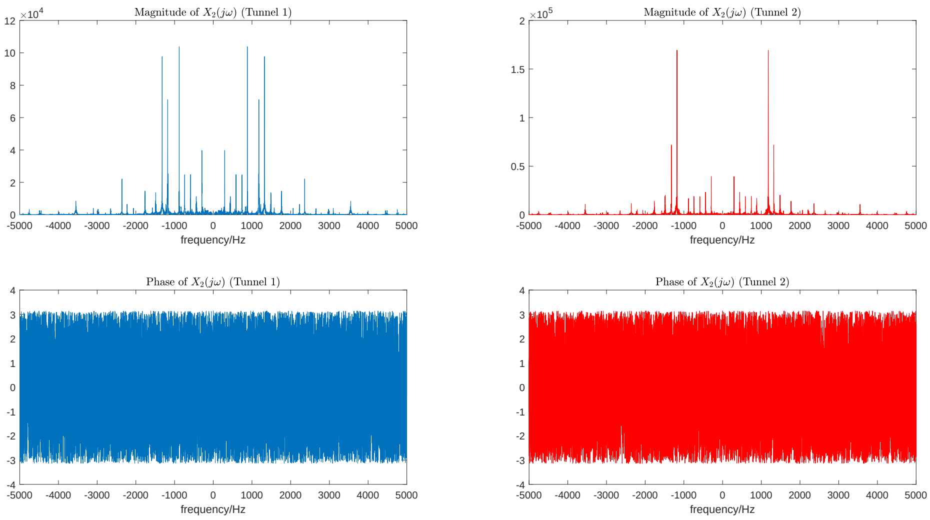


Figure 7: The spectrum of $x(2t)$

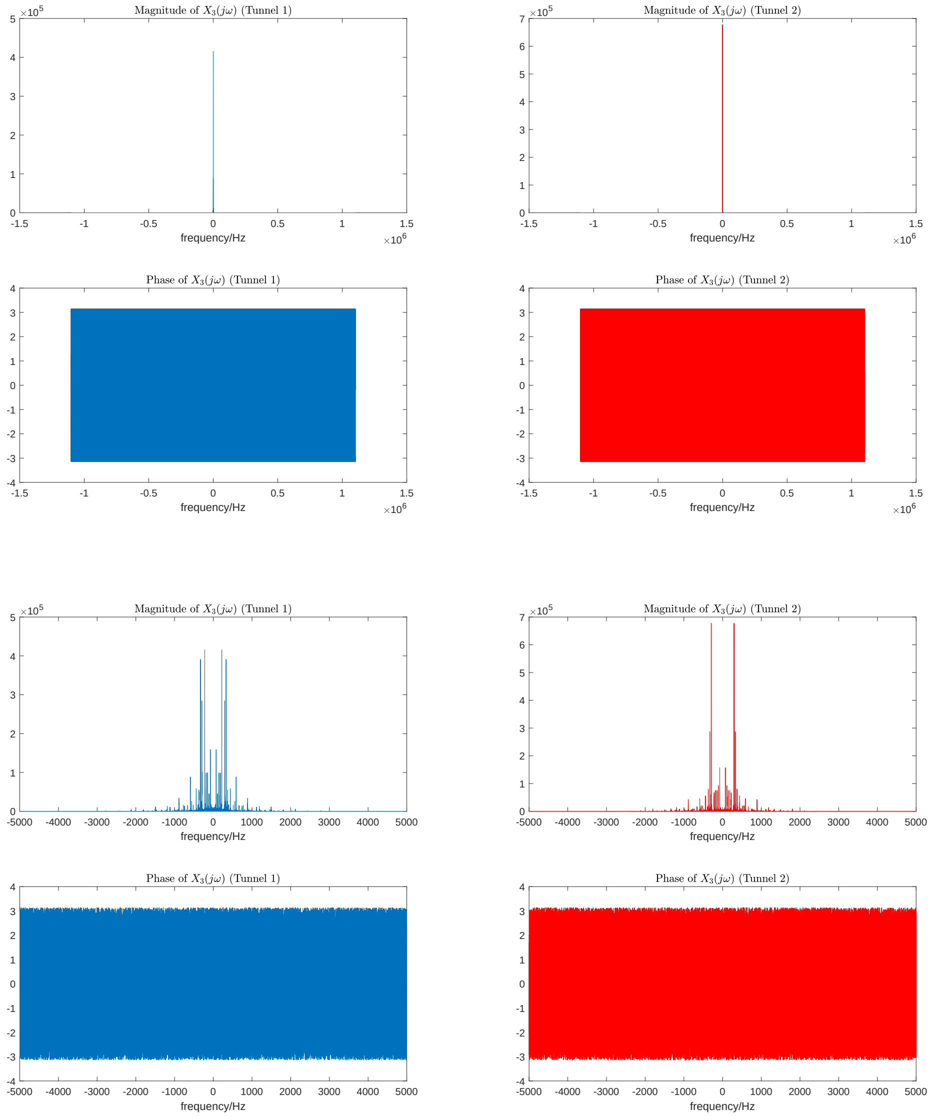


Figure 8: The spectrum of $x(t/2)$

Assumptions and approximations: In this section, I assume that after the linear interpolation, the new signal can be consider as a "continuous" signal, with $f_{sim} >> f_s$, although it's not a really continuous signal. Then we can use the property of continuous time signal to process. I use DFT to approximate CTFT, the computation of the continuous Fourier transform of $x(t)$ can be done approximately using the Fast Fourier Transform (FFT) multiplied by the sampling period [1].

(4). The waveforms reconstructed using only the magnitude spectrum and the phase spectrum, i.e. $\mathcal{F}^{-1}|X(j\omega)|$ and $\mathcal{F}^{-1}e^{jX(j\omega)}$, are shown in Figure 9 and 10.

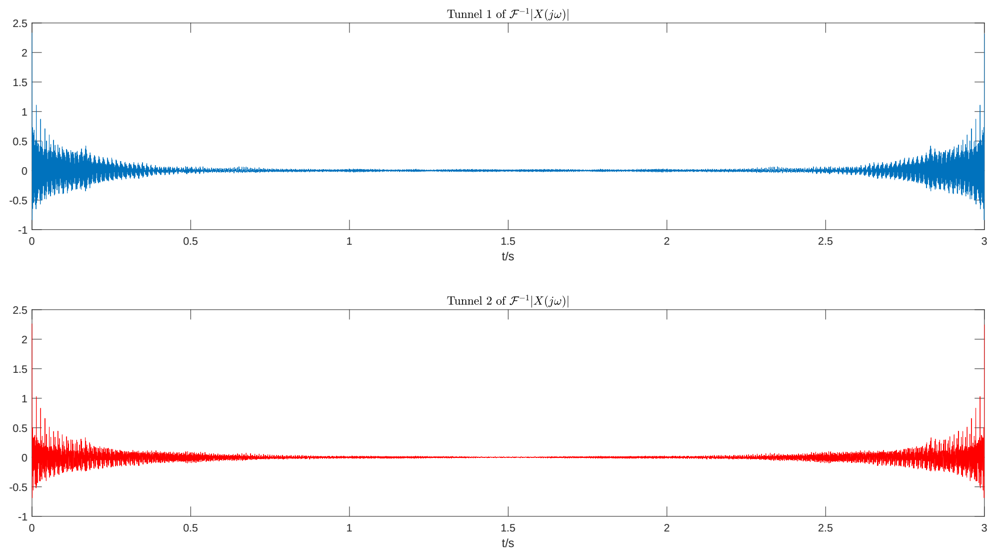


Figure 9: The waveforms reconstructed using only the magnitude spectrum

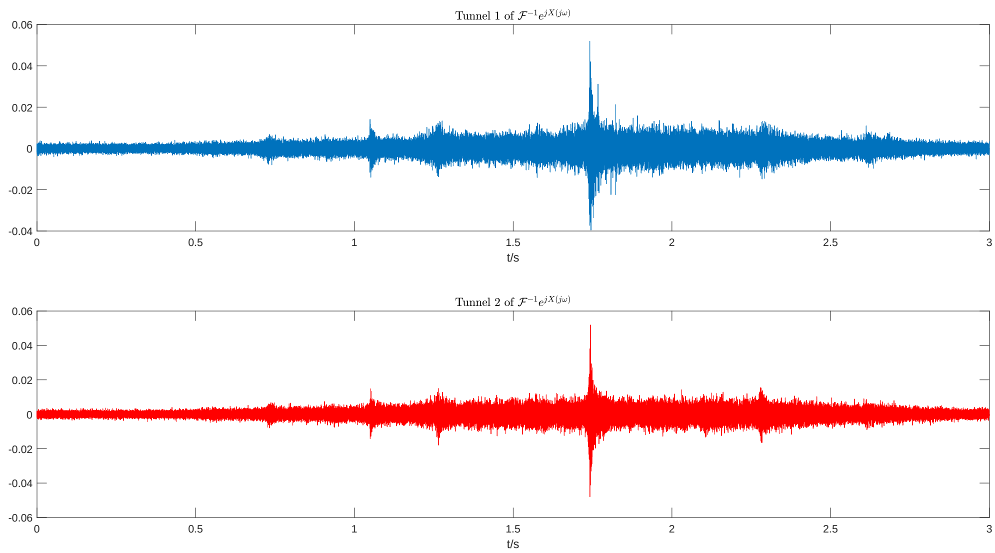


Figure 10: The waveforms reconstructed using only the phase spectrum

Compare them with $x(t)$ in Figure 1, we can see that $\mathcal{F}^{-1}|X(j\omega)|$ is quite different from the original signal, there are only some waveform in the beginning and ending. However, $\mathcal{F}^{-1}e^{jX(j\omega)}$ seems to be more likely to the original signal.

(5). I implement a ideal low pass filter in ILPF.m. I choose $1000Hz$ as the cut-off frequency, then apply it to $x(t)$. The waveform of the output signal of this filter, denoted as $x_{lp}(t)$, is shown in Figure 11. And the spectrum of $x_{lp}(t)$ is shown in Figure 12(a) and 12(b).

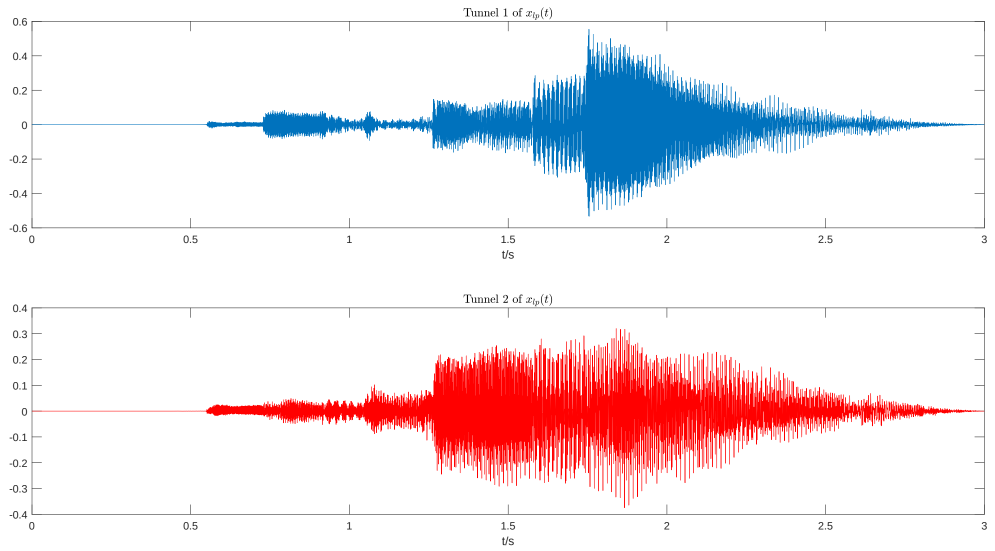


Figure 11: The waveform of the output signal of ILPF

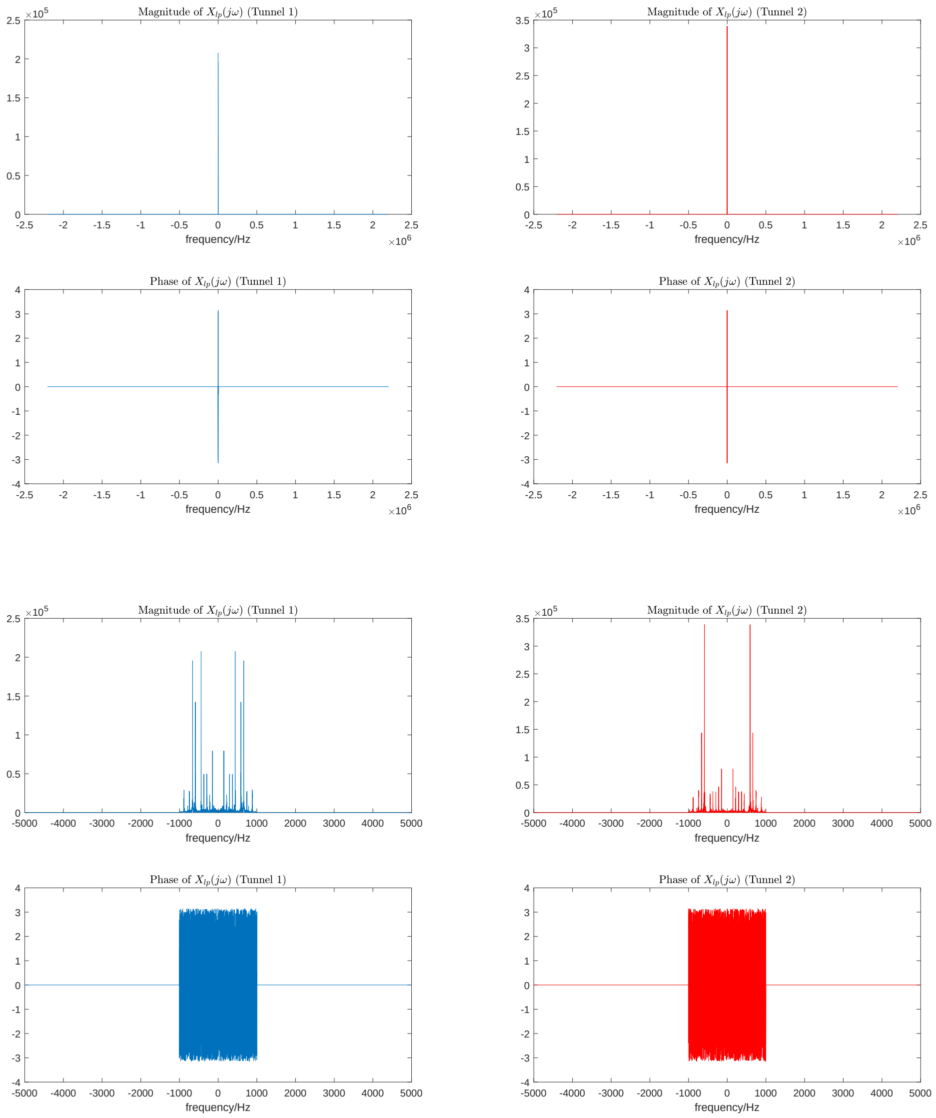


Figure 12: The spectrum of the output signal of ILPF

Part 2: Speech Sampler

(1). In this section, I choose an audio clip and then generate different DT signals by varying the sampling rate. I choose $f_{s,0}$ as the original sampling rate of the audio, which is $44100Hz$. Then I choose the sampling rate as $f_{s,i} = 2^{-i} f_{s,0}$, for $i = 0, 1, 2, 3, 4, 5, 6$. I do the sampling by retain the entries in the signal matrix with index $1 + k * 2^i$, k is a positive integer and must smaller than a fixed upper bound. Then I use `audiowrite` to write each signal to a audio clip. You can hear ta0.wav ta6.wav to find the difference.

After hearing the sampled speech, I find that the audio become more unclear when the sampling rate become smaller. When the sampling rate is $f_{s,4}$, $f_{s,5}$ or $f_{s,6}$, the audio is hard to understand. What's more, when i become bigger, the loss of quality become sharper. In other words, the difference of quality between signals with sampling frequency $f_{s,i}$ and $f_{s,i+1}$ become larger.

(2). In this section, I also use `interp1` to do linear interpolation, which satisfies $f_{sim} >> f_{s,i}$. Then each signal can be viewed as "continuous" time signal. The reconstruction of each DT signal are visualized as follows:

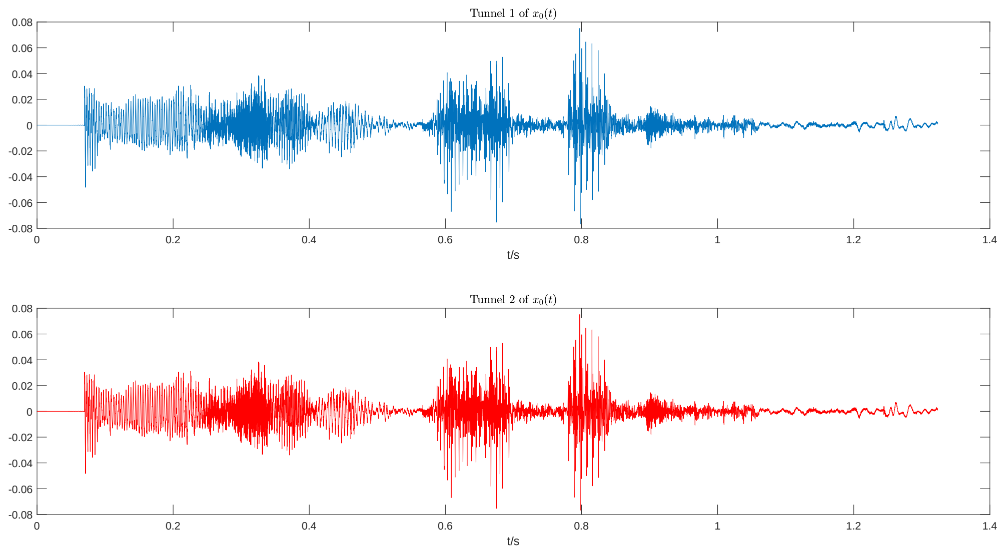


Figure 13: The reconstruction of DT signal with sampling rate $f_{s,0}$

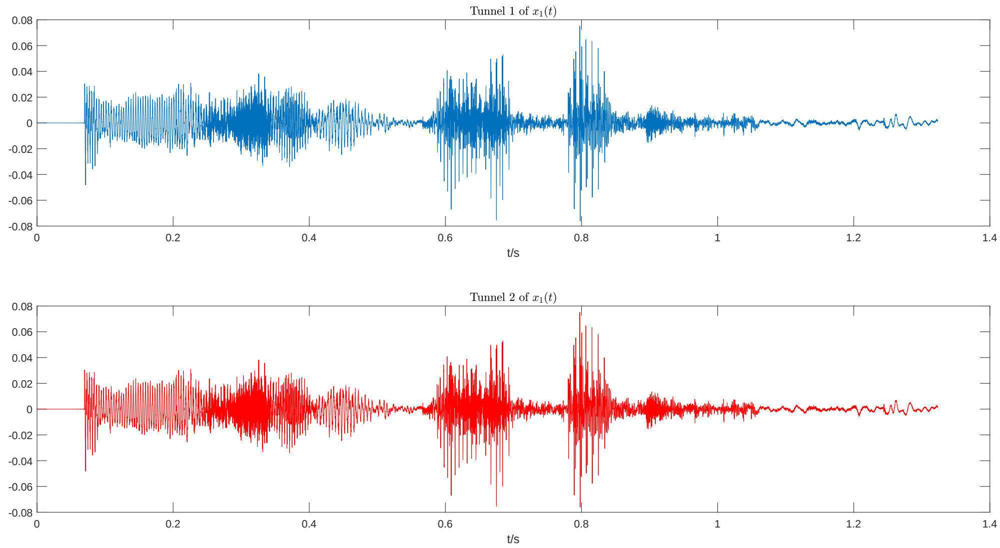


Figure 14: The reconstruction of DT signal with sampling rate $f_{s,1}$

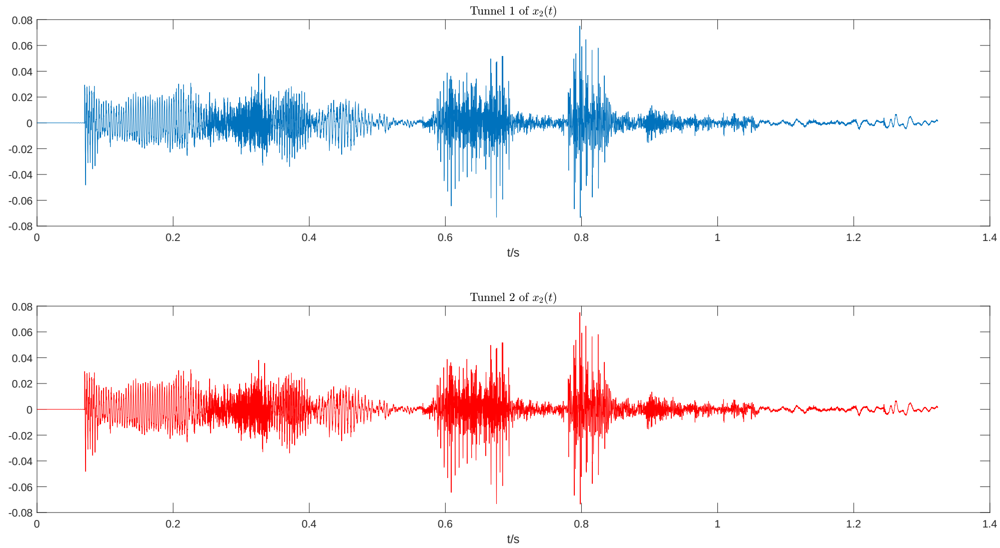


Figure 15: The reconstruction of DT signal with sampling rate $f_{s,2}$

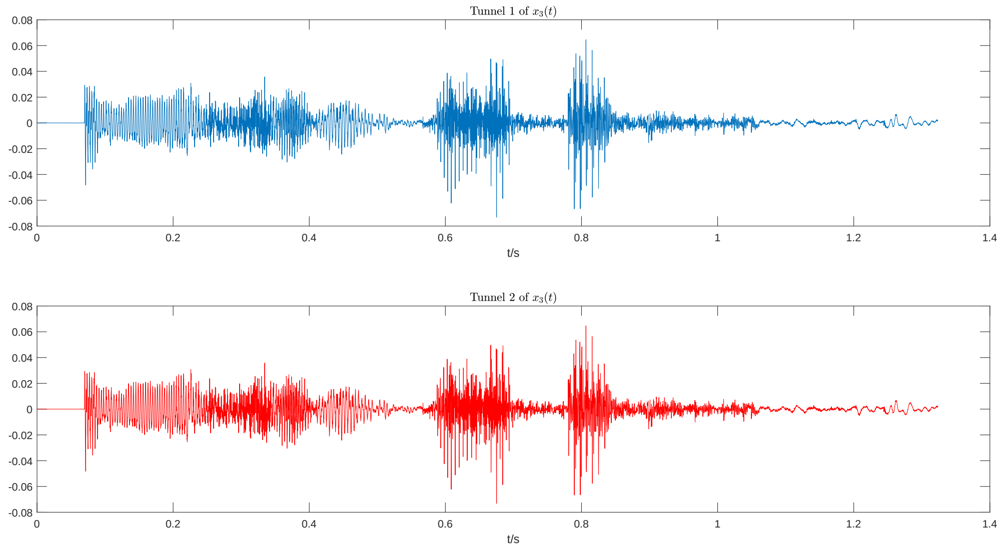


Figure 16: The reconstruction of DT signal with sampling rate $f_{s,3}$

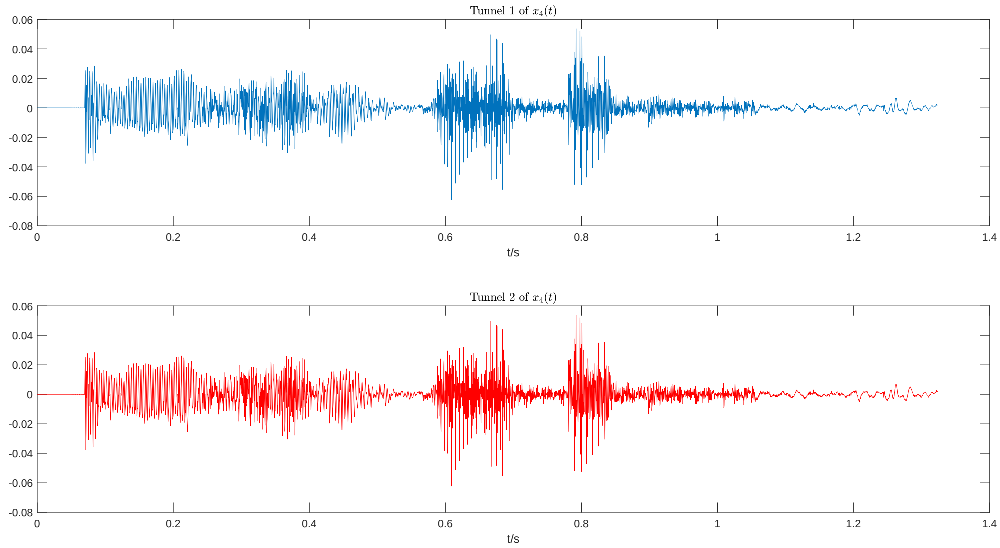


Figure 17: The reconstruction of DT signal with sampling rate $f_{s,4}$

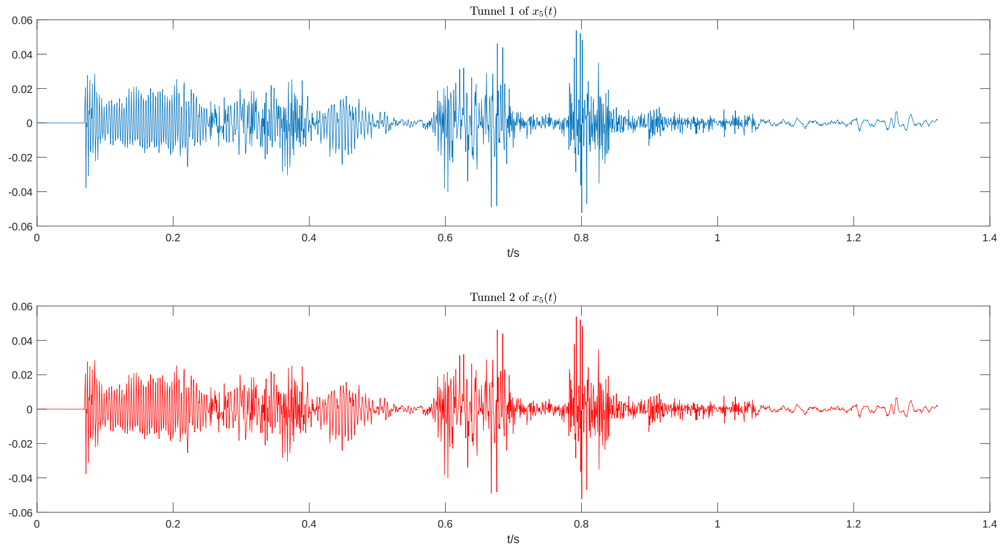


Figure 18: The reconstruction of DT signal with sampling rate $f_{s,5}$

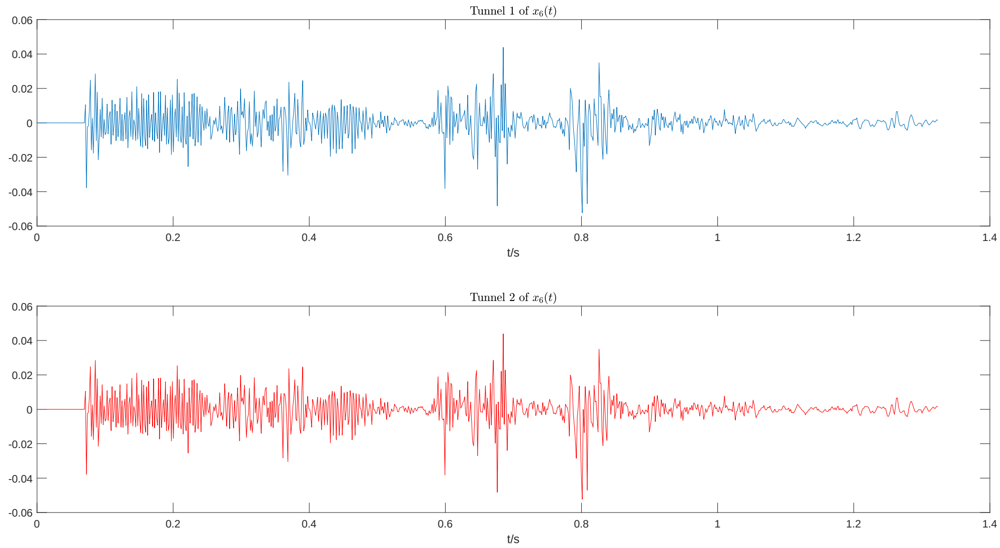


Figure 19: The reconstruction of DT signal with sampling rate $f_{s,6}$

The influence of sampling rate is essential, as the sampling rate become smaller, the influence is deeper. We can first show the spectrum of the original signal $x(t)$.

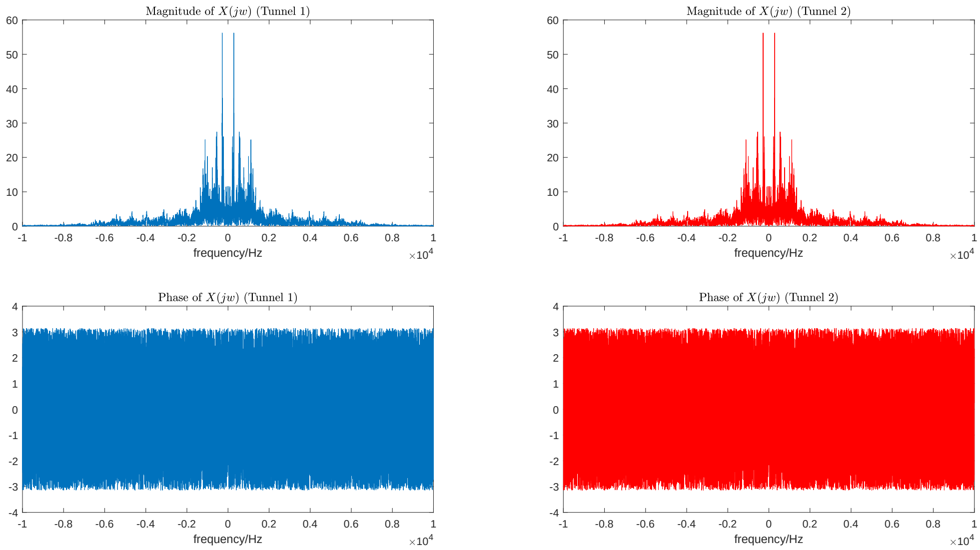


Figure 20: The spectrum of the original signal $x(t)$

We can see that the Nyquist frequency is approximately $8000Hz$, then the Nyquist rate is approximately $16000Hz$. We can get that only $f_{s,0}$ and $f_{s,1}$ are larger than the Nyquist rate. When the sampling rate is smaller, the high frequency components are eliminated more, which causes the reconstruction quality to be worse.

We need a metric to calculate the reconstruction error, my choice is to use the total energy to evaluate the reconstruction error. That's because the sampling rate can affect the components in frequency domain, some high frequency components are lost. So we can use the integral of energy-density spectrum, i.e. $X|(j\omega)|^2$, to denote the completeness of signals. By Parseval's Identity, we can get that the total energy of $x(t)$ satisfies:

$$\|x\|_2^2 = \frac{1}{2\pi}\|X\|_2^2$$

Then we just calculate $\|x\|_2^2$, which can be approximated as the sum of square of "continuous" time signal, i.e. $\sum |x_t[n]|^2$. We suppose $x_0(t)$ is the "perfect" continuous time signal, we calculate the ratio between energy of another signal and the "perfect" one. This ratio is a metric to evaluate the reconstruction error. The smaller the ratio is, the more reconstruction error is to be obtained. The output is:

```
The ratio between x_1(t) and x_0(t) is 0.98 in tunnel 1 and 0.98 in tunnel 2
The ratio between x_2(t) and x_0(t) is 0.94 in tunnel 1 and 0.94 in tunnel 2
The ratio between x_3(t) and x_0(t) is 0.87 in tunnel 1 and 0.87 in tunnel 2
The ratio between x_4(t) and x_0(t) is 0.71 in tunnel 1 and 0.71 in tunnel 2
The ratio between x_5(t) and x_0(t) is 0.63 in tunnel 1 and 0.63 in tunnel 2
The ratio between x_6(t) and x_0(t) is 0.56 in tunnel 1 and 0.56 in tunnel 2
```

Figure 21: The ratio between energy of another signal and the "perfect" one

(3). The basic design idea of the proposed Speech Sampler, as for the MATLAB code, I sample the original DT signal by apply index vector to fetch those entries with index $1 : 2^i : length(signal)$. And then I reconstructe the continuous time signal by linear interpolation.

The priciple of this problem is Nyquist-Shannon Sampling Theorem. Only those sampling rate which is larger than the Nyquist rate of the continuous time signal can be used with no aliasing. This is a special case since the Nyquist frequency is about $8000Hz$, and to specify each word may only need Nyquist frequency to be about $3000Hz$ (In this case, you can see that ta3.wav is still clear enough to understand). So if the requirement is more strict, we can choose the sampling rate to be $22000Hz$ or more, otherwise, we can choose the sampling rate to be $6000Hz$ or more.

Conclusion

In this project, I combined the knowledge in class and the process by code. This project covers many important properties and thoerems. The analysis in time domain and frequency domain can be linked by FFT in practice since signals are always be express as DT signals. And we can use many interpolation method to get a "continuous" time signal.

Acknowledgement

I really want to thank Prof. Jiang, he motivates me to think deeper and find the truth, also helps me to solve many problems. Signals and Systems is really a beautiful subject, from voice to image, from light to wave, it's so close to almost every corner. In the future, I will study Digital Signal Processing, Digital Image Processing, Computer Vision and so on. I thank this course, for it opens a door for me, to search the world.

References

[1] Luis Chaparro and Aydin Akan. *Signals and Systems using MATLAB*. Academic Press, 2018.